Semantic feature based classification of Brain MRI using PCA and PNN

K.Selva Bhuvaneswari.
Assistant Professor
Dept of CSE, University College of Engineering
Kanchipuram, Kanchipuram, India
*bhuvaneswari.ksb@gmail.com

P.Geetha.
Associate Professor
Dept. of CSE, College of Engineering Guindy, Anna University, Chennai-600 025, India.
**geetha@cs.annauniv.edu

Abstract- Semantic based image classification and segmentation is an important but inherently difficult problem in magnetic resonance (MR) medical images. Various image processing techniques are used to detect the abnormalities in the MRI images. Semantics, with respect to images, represents the association between low-level visual features and high-level concepts that can be described in words. An intelligent classification technique is essential to identify normal and abnormal slices of MRI brain images. The proposed method involves the idea of semantic feature layers (SFL) that relates feature classes based on features of lower levels and include additional knowledge. Vital features comprising of statistical information fed by Principal Component Analysis and semantic feature layer is used for classification. The proposed work has two stages: feature extraction and classification. In first stage PCA and SFL is used to achieve the feature extraction. In the next stage, extracted features are fed as input to PNN. It classifies the images as normal and abnormal. The classification results are compared with existing methodologies that does not involve a semantic phase. Performance measures such as sensitivity, specificity and accuracy are used to validate the system.

Keywords: PCA, Semantic feature layer, PNN, Classification.

I. INTRODUCTION

The MRI brain tumor image classification is becoming increasingly important in the medical arena. since it is crucial for surgical planning and intervention. Manual classification of magnetic resonance (MR) brain tumor images is a challenging and time-consuming task. Medical Image Processing has developed to detect as well as diagnose various disorders. The medical images data are acquired from Bio-medical imaging procedures like Computed Tomography scan, Magnetic Resonance Imaging scan and Mammogram scan, which indicates the presence or absence of the lesion. The most important challenge in brain MRI analysis has problems such as noise, intensity non-uniformity (INU), partial volume effect, shape complexity and natural tissue intensity variations[11,9]. Incorporation of a priori medical knowledge is necessary for robust and precise analysis under such conditions. Recent research integrating Image Processing techniques has produced a number of new methods for semantic analysis of image content, which plays a role of great importance, when considering creation and update of image databases and an appropriate delivery of information contained in these databases. Many different methods have been developed for semantic analysis of image utilizing various types of algorithms. However, important semantic information necessary to interpret image content is mostly not represented in single pixels but in meaningful image objects and their mutual relations. These objects may be closely related to fractals or segments, which represent the analyzed image structure units applied for multi-scale image analysis. This work is based on identifying a methodology for image analysis and interpretation of Brain MRI images using the concept of semantics.

The classifications of MRI brain image data as normal and abnormal are important to analysis for the normal patient and to consider only those who have the possibility of having abnormalities[1]. Diagnosis of malignancies can be done automatic with more accuracy in feature extraction and classification of disease. The supervised learning technique such as Probabilistic Neural Network (PNN)[15] is used for classification as it gives better accuracy and performance than other classifiers. Wavelet transform is an alternative tool for feature extraction, because they allows different stages of resolution. This technique requires large storage of data and is computationally more expensive[3]. Hence an alternative method for dimension is to reduce large amount of data scheme is used and increase the discriminative power. So the principal component analysis (PCA) has been used[5]. Principal component analysis is appealing since it effectively reduces the dimensionality of the data and therefore decrease the computational cost of analyzing new data. [8,10] Additionally semantic feature is also used for classification[12].
II. METHODOLOGY

The proposed approach takes the similar MRI images and its feature is extracted for training the database using PCA & SFL technique. The main idea behind using PCA in this approach is to reduce the large dimensionality of data. This leads to more constant and accurate classifier. The high dimensionality of data does not produce graphical representation. PCA[6] is a powerful tool for analyzing data. The main advantage of PCA is that once the patterns are found they can be reduced without much loss of information. Later semantic features are extracted using the concept of Semantic Feature Layer(SFL). The overall block diagram is shown below as,

![Proposed methodology](image)

III. FEATURE EXTRACTION

Principal Components Analysis (PCA) is a method for detecting features in data, and expressing the data in such a way as to highlight their similarities and differences. Since patterns in data can be difficult to find in(segmentations) and combine them with an ensemble function we propose to obtain multiple object partitions. It is used as a feature extraction algorithm. The principal component analysis (PCA) is one of the most successful techniques that have been used in image recognition and compression.[11] The objective of PCA is to reduce the dimensionality of the large number of data. Image compression is based on compress the data by reducing large number of dimension without loss of information.

A. Algorithm 1 PCA algorithm

Let X be an input data set (X: matrix of dimensions M X N).

Perform the following steps:

1: Calculate the empirical mean u[m] = (1/N) PN i=1 X[m,n].

2: Calculate the deviations from the mean and store the data in the matrix B[M N]:, B=X-u.h, where h is a 1 x N row vector of all 1’s: h[n] = 1 for n=1....N.

3: Find the covariance matrix C: C = (1/N)B.B^∗.

4: Find the eigenvectors and eigenvalues of the covariance matrix V −1CV = D, where V: the eigenvectors matrix, D: the diagonal matrix of eigenvalues of C. D[p,q] = λm for p=q=m is the mth eigenvalues of the covariance matrix C.

5: Rearrange the eigenvectors and eigenvalues λ1 ≥ λ2 ≥ λ3 ≥ ...... ≥ λN.

6: Choosing components and forming a feature vector. Save the first L columns of V as the M x L matrix W: W[p,q]=V[p,q] for n=1....M, q=1....L where 1 ≤ L ≤ M.

7: Deriving the new data set The eigenvectors with the highest eigenvalues are projected into space, this projection results in a vector represented by fewer dimension (L < M) containing the essential coefficients.

This algorithm is used to calculate the principal components of the input image.

B. Semantic Features Extraction

The idea of semantic feature layers (SFL) is the design of semantically related feature classes that are based on features of lower levels and include additional knowledge. Additional knowledge can be comprised of modeling information, domain knowledge, statistical information, etc. and be expressed as data (e.g. a color covariance matrix) or as algorithms[3,4] (e.g. a sophisticated distance measurement algorithm). SFL should help to reduce the size of the semantic gap. SFL are more than DS. Descriptor Schemes (DS) are containers of Descriptors (D) and DS. DS define hierarchical relationships of static
Descriptors and other DS. In SFL, Descriptors do not remain static on higher levels but are transformed by additional knowledge to more specific [12](semantic) representations.

Using SFL in addition or instead of low-level features has two major advantages:

- It is possible – in the context of the SFL – to perform high-level queries without the need to translate them to queries on low-level features. This should lead to better results.
- Queries are much faster, because of simpler feature vectors and simpler querying methods. The integration of additional knowledge on the basis of low-level features will in most cases lead to a compression of the high-level feature vectors. This process is performed offline during the feature extraction process. Querying methods can be simpler because no mapping is necessary and feature vectors are simpler.

**a) Low-level image features**

Many sophisticated feature extraction algorithms have been designed and good surveys are available [4]. Some of the low-level image features are

- Color
- Texture
- Shape
- Spatial Location

**Color feature**

Color feature is one of the most widely used features in image retrieval. Colors are defined on a selected color space. Variety of color spaces are available, they often serve for different applications. Color spaces shown to be closer to human perception and used widely are RGB, LAB, LUV, HSV (HSL), YCrCb and the hue-min-max-difference (HMMD). Most of those color features though efficient in describing colors, are not directly related to high-level semantics. For convenient mapping of region color to high-level semantic color names, some systems use the average color of all pixels in a region as its color feature.

**Texture feature**

Texture is not so well-defined as color features, some systems do not use texture features. However, texture provides important information in image classification as it describes the content of many real-world images such as fruit skin, clouds, trees, bricks, and fabric. Hence, texture is an important feature in defining high-level semantics for image retrieval purpose. In some systems, texture features are obtained based on the texture property of pixels or small blocks contained in the region. For example, for each region, the mean of the texture features of all the 4 *4 blocks it contains is used as the region feature. The problem of such feature is that they cannot sufficiently describe the texture property of the entire region. An intuitive way to solve this problem is to extend the arbitrary-shaped region into a rectangular area by padding some values outside the boundary and then apply block transforms. However, as regions in real-world images are usually not homogeneous texture, such initial padding will introduce spurious components which do not describe the original region and hence degrade the performance of the texture feature obtained. Still another possible solution is to obtain an inner rectangle (IR) from a region onto which block transforms can be performed to generate coefficients from which texture feature can be calculated. This works well when the region texture is homogeneous and the IR carries enough information to describe the region’s texture property. However, image regions in real-world images are usually not homogeneous.

**Shape feature**

Shape is a fairly well-defined concept. Shape features of general applicability include aspect ratio, circularity, Fourier descriptors, moment invariants, consecutive boundary segments, etc. Shape features are important image features though they have not been widely used as color and texture features[3]. Shape features have shown to be useful in many do-main specific images such as man-made objects.

**Spatial Location feature**

Besides color and texture, spatial location is also useful in region classification. For example, ‘sky’ and ‘sea’ could have similar color and texture features, but their spatial locations are different with sky usually appears at the top of an image, while sea at the bottom. Spatial locations usually are simply defined as ‘upper, bottom, top’ according to the location of the region in an image. Relative spatial relationship is more important than absolute spatial location in deriving semantic features. 2D-string and its variants are the most common structure used to represent directional relationships between objects such as ‘left/right’, ‘below/above’. However, such directional
relationships alone are not sufficient to represent the semantic content of images ignoring the topological relationships.

Keyword indexing techniques can be used to capture an image’s semantic content, describing objects clearly identifiable by linguistic cues. These techniques assign keywords or classification codes to each image when it is first added to the collection and use these descriptors as retrieval keys at search time. These kinds of techniques are often encountered in both newspaper and art libraries. Their advantages consist of high expressive power and the ability to describe image content from the primitive level (low-level features) to the abstract level (high-level features); the high level features involve a significant amount of reasoning about the meaning and purpose of the objects or scenes depicted (such as subjective emotions associated with an image). One of the drawbacks of the current manual indexing techniques is the time of assigning the keywords. When the indexing time for every image takes several minutes, the indexing for a considerable large collection of images is an intensive and time consuming task. Beside the amount of effort that is needed in order to complete such a job, manual indexing has another drawback that cannot be solved with either time or labor. This drawback comes from the fact that the same picture can have different meanings for different people or even for the same person at different times. On the other hand, there are images (such as trademarks) that cannot be described by linguistic cues. Therefore, more effectively indexing techniques are necessary.

b) Semantic feature extraction with shape as low level feature

For extracting the semantic features the following steps are proceeded to map the low-level features with the high-level semantic features.

1. Shape as Low-level feature
2. Keyword as High-level feature
3. Interpretation of low-level and high-level features

1. Shape as Low-level Feature

In this paper, one of the low-level features named as, shape feature is used, to map with the semantic information of the image. In this paper, the edge detection is performed using Slope Magnitude method. Shape feature extraction requires the edges of the extracted shape to be connected in order to reflect the boundaries of objects present in the image[4]. For extracting shape features as the form of connected boundaries, from the given MRI Brain Image, Sobel Gradient Operator is utilized.

Sobel Gradient Operator

It is a discrete differentiation operator, that is used to performing an approximation of the gradient of the image intensity function. Convolving the image with a small, separable and integer valued filter in horizontal and vertical direction, is the operation performing by the Sobel Operator. Based on computations, Sobel Operator is inexpensive.

Steps to apply Slope Magnitude method

Step 1 : Convolve the original image with the Sobel mask $S_x$ and $S_y$. $S_x$ mask is used to obtain the $x$ gradient and $S_y$ mask is used to obtain the $y$ gradient.

$$S_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} \tag{1}$$

$$S_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \tag{2}$$

Step 2 : Get the individual squares of both values as $S_x^2$ and $S_y^2$ in equation (1) and (2).

Step 3 : Add the two squared terms as $S_x^2 + S_y^2$.

Step 4 : Take square root of the sum and then we get the equation as,

$$S = \sqrt{S_x^2 + S_y^2} \tag{3}$$

Thus low-level shape feature is obtained using the slope magnitude method.

2. Keyword as High-level Feature

Keywords are the features that are helpful for describing the high-level domain concepts. The attributes of semantics includes some subjectivity, uncertainty etc. The shape edge property is directly extracted using edge detection. High level semantic property can be extracted as the keyword, on the basis of low-level visual feature, shape. In the MRI Brain Image, the shape edge is initially extracted and according to the clearance of shape edge. The semantic terms that are related with the clearance of the shape edge are:

$$ST = \{"low", medium, high\}$$

3. Interpretation of low-level and high-level features:

Initially find out the bounds of the semantic class, clearance.

(i. e.) "low = $\alpha_1$, medium = $\alpha_2$, high = $\alpha_3$"

$$ST = \{"low = \alpha_1$, medium = $\alpha_2$, high = $\alpha_3\} \tag{4}$$

For mapping the shape feature edge $S$ into semantic term $ST$, the following inference rules or the degree of clearance are used.
The lower inf and upper sup bounds are representation of shape feature edge and both are related to $\alpha_1$ and $\alpha_3$, respectively.

Thus, the statistical features using PCA and semantic features using SFL are extracted for further classification. The extracted features are given to the Probabilistic neural network[16]. PNN will train the neural network with the Eigen values extracted from the MRI and the semantic value $\alpha$. Then the test sample is compared with the extracted features and the abnormalities are identified.

IV. CLASSIFICATION USING PROBABILISTIC NEURAL NETWORK

Probabilistic Neural Network (PNN) and General Regression Neural Networks (GRNN) have resemblance architectures, but there is a fundamental difference.[15] Probabilistic neural networks perform classification where the output variable is categorical, whereas general regression neural networks perform regression where the output variable is continuous.

The standard PNN structure have four layers:

A. INPUT LAYER

It has one neuron in the input layer for each predictor variable. The categorical variables has N-1 neurons where N is the number of categories. It calculates the range of the values by subtracting the median and dividing by the interquartile range. The input neurons are then fed the values to each neurons in the hidden layer.

B. HIDDEN LAYER

This layer has one neuron for each case in the training data. Each neuron stores the values to the predictor variables for the case along with the output value. The vector values from the input layer, a hidden neuron compute the Euclidean distance of the test case from the neuron and also applies the RBF kernel function using the sigma value(s). Then final value is passed to the neurons in the pattern layer. A description of the PNN classifier was given. PNNs had been used for training and classification purposes. The PNN classifier presented good accuracy, very small training time. The MR images pixels are converted into matrices format by using MATLAB. Finally, PNN is used to classify the MR images.

C. PATTERN/SUMMATION LAYER

The pattern layer in the neural network is different for PNN networks and for GRNN networks. For PNN networks has one pattern neuron for each category of the target variable. The target variable of each training case is stored with each hidden neuron. The weighted value is coming out of the hidden layer[14]. For GRNN networks has only two neurons in the pattern layer. First neuron is the numerator summation unit sums of the weight values multiply by the actual target value for each hidden neuron. other neuron is the denominator summation unit sums of the weight values coming from each of the hidden neurons.

D. DECISION LAYER

The decision layer is different for PNN and GRNN networks[13]. For PNN decision layer comparing the weighted values for each target value gathered in the pattern layer and uses the largest value to predict the target value. For GRNN decision layer dividing the value collected in the numerator summation unit by the value gathered in the denominator summation unit and finally uses the result as the predicted target value.

The abnormal image is classified as benign, malignant. Then tumor is detected by using pnn classification. Finally tumor is segmented as multilevel graph partitioning algorithm.
V. PERFORMANCE EVALUATION

The performance of the proposed algorithm to classify the images obtained using the proposed methodology is compared with the corresponding ground truth images. The proposed technique is analyzed with the following quality parameters are used to study its performance:

- Sensitivity \( [Se = \frac{T_{pos}}{T_{pos} + F_{neg}}] \)
- Specificity \( [Sp = \frac{T_{neg}}{T_{neg} + F_{pos}}] \)
- Positive predictive value \( [Ppv = \frac{T_{pos}}{T_{pos} + F_{pos}}] \)
- Negative predictive value \( [Npv = \frac{T_{neg}}{T_{neg} + F_{neg}}] \)
- Accuracy \( [Acc = \frac{T_{pos} + T_{neg}}{T_{pos} + F_{neg} + T_{neg} + F_{pos}}] \)

where, \( T_{pos} \) is True positive, \( T_{neg} \) is True negative, \( F_{pos} \) is False positive and \( F_{neg} \) is False negative. The parameters, \( Se \) and \( Sp \) define the ratio of well classified normal and abnormal images, respectively. \( Ppv \) is the ratio of images classified as normal that have been correctly classified. \( Npv \) is the ratio of images classified as abnormal that are correctly classified. Lastly, \( Acc \) is the ratio of total well detected and classified normal images. All these parameters help in defining the performance of the proposed technique as explained in the previous sections and are tabulated in Table 1.

The performance of our proposed algorithm using is tabulated in Table 2. The same is graphically illustrated in Fig 4. The experimental results prove that the accuracy rate and sensitivity of proposed methodology to be higher compared with other conventional methodologies.

<table>
<thead>
<tr>
<th>Images</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.99</td>
<td>0.99</td>
<td>0.97</td>
</tr>
<tr>
<td>2</td>
<td>0.98</td>
<td>0.99</td>
<td>0.98</td>
</tr>
<tr>
<td>3</td>
<td>0.987</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td>4</td>
<td>0.96</td>
<td>0.97</td>
<td>0.96</td>
</tr>
<tr>
<td>5</td>
<td>0.97</td>
<td>0.98</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Table 1. Performance evaluation of the semantic based Classification method

<table>
<thead>
<tr>
<th></th>
<th>KNN</th>
<th>FNN</th>
<th>PNN&amp;SFL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>0.86</td>
<td>0.91</td>
<td>0.99</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.56</td>
<td>0.97</td>
<td>0.995</td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.89</td>
<td>0.88</td>
<td>0.995</td>
</tr>
</tbody>
</table>

Table 2. Performance comparison of the proposed method with conventional methods

VI. CONCLUSION

In this work, a new approach for the semantic based classification of MRI brain images is proposed. This comprises of automatically recognizing semantically meaningful regions in an image. This work aims to correlate each pixel in the image with a label denoting a semantically meaningful part. It helps the physician and radiologist for further malignancy detection and diagnosis. The local binary patterns and gray level co-occurrence features are extracted from brain images with benign and brain images with malignant and normal brain images. These extracted features along with semantic feature are trained using PNN classifier in training mode. The same features are extracted from test brain image and classified with trained patterns using PNN classifier in classification mode. This proposed computer aided automation system for classification achieves 99.4% of sensitivity, 99.6% of specificity, 97.03% of positive predictive value and 99.5% of overall accuracy. The results show that the Semantic approach is abstractly dissimilar from classical methodology with the final aim being the correct category association to image areas rather than existing methodologies.

References


[7]. Samir Kumar Bandhyopadhyay, Tuhin Utsab Paul, “Automatic Segmentation of Brain Tumor from Multiple Images ofMRI Brain”, International Journal of Application or Innovation in Engineering & Management (IJAIME) Volume 2, Issue 1, January 2013 ISSN 2319 – 4847


[13]. Xiaoli Zhang Fern, Carla E. Brodley brodley, “Solving Cluster Ensemble Problems by Bipartite Graph Partitioning” School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907